

Il genere nell'intelligenza artificiale

Silvana Badaloni

Dipartimento di Ingegneria dell'Informazione,

Centro di Ateneo Elena Cornaro

Università di Padova

silvana.badaloni@unipd.it



S. Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 16 maggio 2025

Genere e Scienza

THREE FIXES

1. FIX THE NUMBERS OF WOMEN
2. FIX THE INSTITUTIONS
3. FIX THE KNOWLEDGE

Fix the numbers of women

<http://www.unipd.it/bilancio-genere>

*The world need Science and Science needs Women.
And good solutions need reliable statistics and indicators*



Studentesse immatricolate Ingegneria

Ing. IT

- 26,3 % Proporzione di donne immatricolate nei corsi ingegneristici 21/22 (25,2 % 19/20);

Ing. UNIPD

- Valor medio **24,3 %** AA 24/25
- Ingegneria per l'Ambiente e il Territorio 30,3 %
- Ingegneria Gestionale 30,2 %
- Ingegneria Civile 29,8 %
- Ingegneria Meccanica 11,8 %

DEI

- aa 24/25 Valor medio **25,3 %**
- Ingegneria biomedica 57,8 %
- Elettronica 9,7 % - Ingegneria Informatica 11,8 % (15,4 % 23/24)

Sguardo al settore ICT

Information and Communication Technologies

- In 35 paesi europei meno di 1 su 5 laureati in computer science sono donne¹
- Divario di genere in un settore in grande sviluppo
- Rapporto di Almalaurea 2022² (Consorzio interuniversitario) ci dice che:
 - i/le laureati/e STEM si posizionano meglio rispetto al totale in tutte le discipline
 - sono le laureate in Ingegneria Industriale e dell'Informazione quelle con le retribuzioni mensili più elevate (1.723 euro) sebbene inferiori a quelle dei colleghi maschi (1.872 euro) - gender pay-gap
 - si tratta di una evidenza che va divulgata in ambito scolastico per aiutare le giovani a fare scelte consapevoli

1. OECD Gender Data Portal, Where are tomorrow's female scientists?

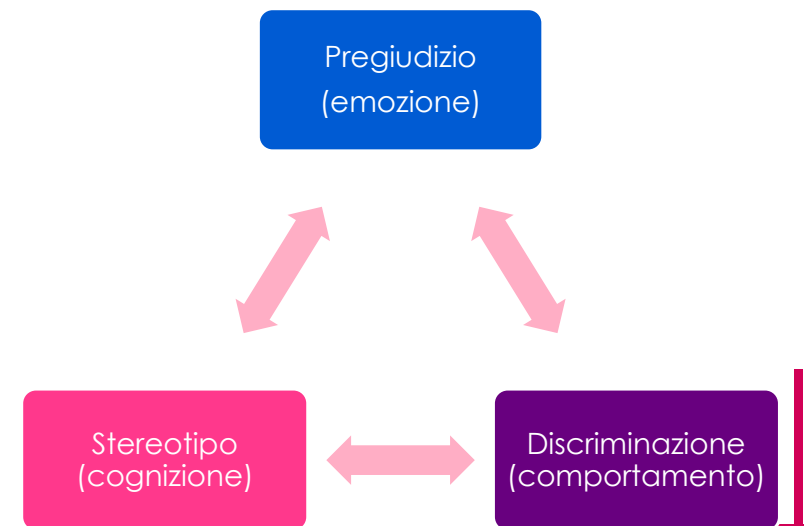
2. Rapporto Almalaurea 2022. Laureate e laureati: scelte, esperienze e realizzazioni professionali

https://www.almalaurea.it/sites/almalaurea.it/files/convegni/gennaio2022/6_almalaurea_rapportocompleto_laureatelaureati.pdf

Bias


- **Stereotipo o bias** è un sistema di conoscenza condivisa nella società, pro o contro qualcosa
- Gli stereotipi ci aiutano a confrontarci con un mondo diverso/sconosciuto, assumono allora lo status di giudizi che categorizzano fatti o persone: pregiudizi
- **Pregiudizio** è un'opinione preconcepita concepita non per conoscenza precisa e diretta del fatto o della persona, ma sulla base di voci e opinioni comuni
- **Circolo vizioso**

(source: F.A. Lisi)



Ruolo degli stereotipi

- Ruolo degli stereotipi che condizionano la scelta delle giovani e dei giovani
- Stereotipo sulla Scienza: ce n'è uno che porta a considerare una **fatalità** il fatto che ci siano ancora poche donne nella Scienza
- Uno molto diffuso riguarda la presunta incapacità delle donne in ambito logico-matematico
- Non è vero che la presunta inferiorità femminile sia dovuta ad una diversità a livello cerebrale tra i due sessi: studi scientifici hanno mostrato che questa idea è falsa, ed è un pregiudizio¹
- Il cervello non ha sesso (Antonella Viola)
- E' un problema culturale che va affrontato, a tutti i livelli in quanto il contesto familiare di provenienza, i modelli sociali proposti, l'attitudine delle scuole hanno una forte influenza sulle scelte formative delle giovani generazioni



J.S.J.Ceci, W.M.Williams, S.M.Barnett. Women's underrepresentation in Science: sociocultural and biological considerations. Psychological Bulletin, 2009, vol 135, 2, 218-261

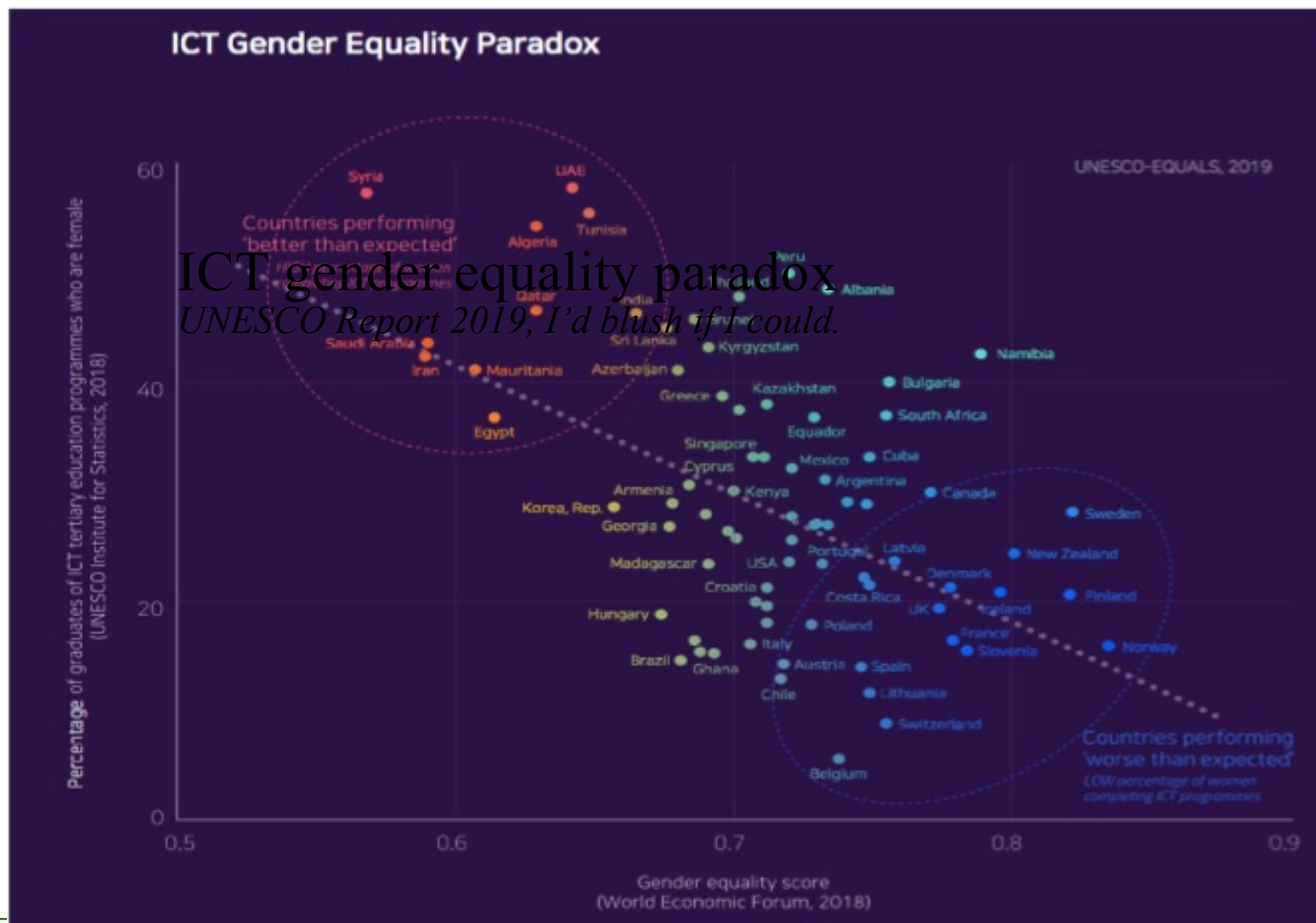
Quale ingegneria?

- Le ragazze, se scelgono ingegneria, sono selettive e scelgono corsi nuovi e interdisciplinari. I numeri delle giovani iscritte sembrano dipendere da quanto distante dalla società è percepito il sapere, la conoscenza nell'ambito dello specifico corso di ingegneria
- Le ragazze non si identificano nell'immagine stereotipata dell'ingegnere: rigido, pesante alleato delle macchine e della tecnologia
- Ingegneria/tecnologia in un contesto sociale/delle scienze della vita
- NERD



ICT gender equality paradox

UNESCO Report 2019, I'd blush if I could.



Donne ed AI



- In Europa e nel Regno Unito solo il 16% delle persone che lavorano nel campo IA sono donne e solo il 12% ha più di 10 anni di esperienza. Questo dato si ripresenta anche in Italia dove le donne ricoprono solo il 16% dei posti di lavoro dell'IA
- Le istituzioni e le aziende devono aumentare non solo il numero delle donne tech, ma anche la loro rappresentanza nei ruoli apicali, un maggior numero di donne in posizioni dirigenziali può sostenere politiche a favore della salute delle donne
- Superare il tetto di cristallo



Fix the Institutions

Progetti Europei



S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 16 maggio 2025

Structural Changes Projects

- La strategia della Commissione Europea è stata, ed è tuttora, quella di promuovere cambiamenti strutturali nelle Istituzioni di ricerca grazie a politiche di genere specifiche.
- Dal 2011 la EC ha finanziato molti EU Projects: GenisLAB, INTEGER, STAGES, FESTA, **GenderTIME** (<https://gendertime.org>), GENOVATE (<http://www.genovate.unina.it/>), TRIGGER, EGERA, GARCIA, **GENERA**(<https://genera-project.com/>), PLOTINA(<http://www.plotina.eu/>), EQUAL-IST, ect
- Molte Università e centri di ricerca hanno adottato tailored **Gender Equality Plans (GEP)** come un diretto risultato del loro coinvolgimento in EU projects per promuovere cambiamenti strutturali

Ruolo della Politica EU è stato rilevante

FP7 EU GenderTIME

Trasferring Implementing Monitoring Equality

- 10 partners in Europa - coordinatrice Yvonne Pourrat (Ecepie, France)
- L' Università di Padova era partner – coordinatrice scientifica del UNIPD team: Silvana Badaloni
- 2013-2016
- www.gendertime.org
- *The aim of the GenderTIME project was to implement self-tailored action plans to increase the participation and career advancement of women researchers*
- Come si può misurare la Parità di genere: indicatore UNIPD-GEI
- Obbligo attuale di GEP per progetti di ricerca in tutti i campi del sapere: requisito per l'accesso ai finanziamenti del programma di ricerca Horizon



Fix the knowledge Innovazioni di genere

- Genere e scienza non vuole dire solo aumentare il numero delle donne presenti per battere la sotto-rappresentazione delle donne attraverso adeguate politiche di genere
- *Fix the knowledge* vuol dire integrare la dimensione di genere nei contenuti scientifici dell'innovazione.
- Come possiamo ri-disegnare le teorie scientifiche tenendo conto della dimensione di genere?
- According to (Sanchez, 2013): *“There is a need to go beyond stereotypical feminization of products – so called “pinking” – as female preferences can be drivers for substantial innovation”*
- Non e' sufficiente applicare il cosiddetto “pinking” method – i.e. la *stereotipica femmininizzazione del prodotto*



Come?


- Come possiamo dare una nuova interpretazione dei fatti rispetto a un universo proposto come neutro (ma neutro non è)?
- Per produrre una scienza e una innovazione di genere, è necessario radicalmente cambiare la teoria e porre nuove domande scientifiche, riformulare le domande scientifiche, in qualunque campo del sapere
- Usando le regole dell'argomentazione logica abbiamo sviluppato una riflessione formale sul metodo scientifico e sulle modalità per produrre innovazione di genere [Badaloni, Lisi]

Argomentazione falsificante

$H \rightarrow O$

$\neg O$

$\neg H$



quando le conseguenze di una teoria non sono verificate nel contesto sperimentale allora la teoria deve essere completamente ri-progettata

Gendered Innovation

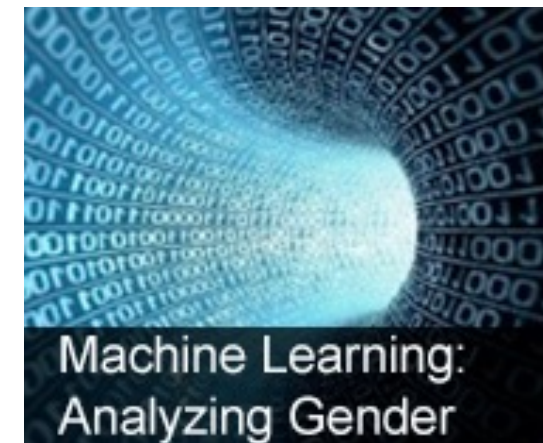
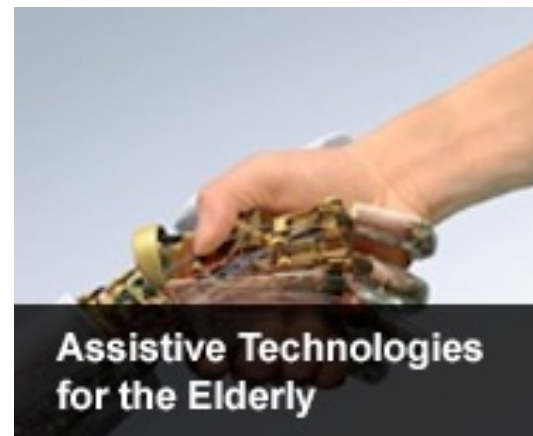


- Londa Schiebinger – Stanford University
- <http://genderedinnovations.stanford.edu/index.html>
- Perché questo approccio?
- Esempio iconico. Negli Stati Uniti, tra il 1997 e il 2000, sono stati ritirati 10 farmaci a causa degli effetti dannosi sulla salute (alcuni potenzialmente mortali) e 8 su 10 si sono rilevati più dannosi per la salute delle donne rispetto agli uomini
- L'approccio va rovesciato: l'innovazione di genere arricchisce la scienza e apre nuovi orizzonti. Solo così si può produrre una ricerca che sia gender-responsible, che permetta una crescita della qualità della vita in modo paritario per tutte e per tutti.



Londa Schiebinger, Stanford University

Case studies



<http://genderedinnovations.stanford.edu/index.html>


S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 16 maggio 2025

Heart diseases in Women

- La cardiopatia ischemica è il killer numero uno della popolazione statunitense ed europea, indipendentemente dal genere. Era stata considerata una malattia prevalentemente maschile, nelle donne spesso non era stata diagnosticata
- Gendered innovation:
 1. Ridefinire la patologia (diversa tra donne e uomini)
 2. Sviluppare nuove tecniche diagnostiche
 3. Capire le differenze dovute al sesso nei sintomi
 4. Fattori di rischio (fumo)
 5.



<https://genderedinnovations.stanford.edu/case-studies/heart.html>

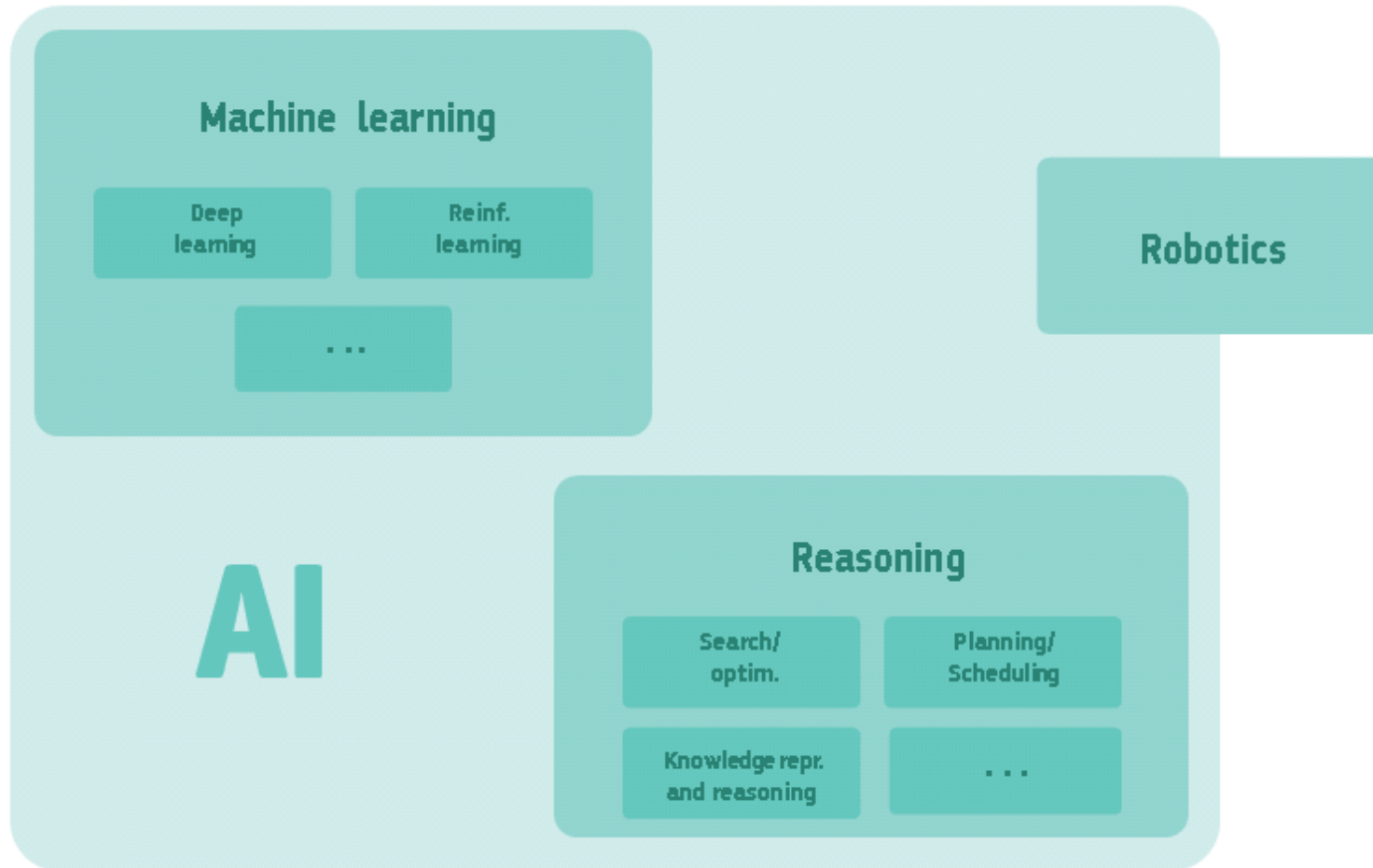


**Heart Disease in
Women: Formulating
Research Questions**

Decliniamo la **Innovazione di genere nell'ambito dell'AI**: quali sono i problemi che il sempre maggiore diffondersi di tali applicazioni possono implicare a livello etico? Quale è il rapporto con le questioni di genere per poter parlare di gendered innovation in questo ambito?
Gli algoritmi possono essere sessisti o razzisti?



General architecture



AI generativa

- AI Generativa: poggia il suo funzionamento sul **modello linguistico**. Nessuna semantica
- LLMs Sistemi addestrati per predire la probabilità della prossima parola (***next word prediction***) dato il contesto attorno a quella parola^[1]_{SEP}
- L'AI non è certo nata un anno fa, con il lancio di **ChatGPT**. Perché allora sembra che ci sia come un'era pre e post Gpt (novembre 2022)? Siamo davvero davanti a qualcosa che cambierà il mercato e tutti i soggetti?



La nascita dell'AI

- Conferenza di Dartmouth 1956
 - John McCarty ha coniato il termine **Artificial Intelligence**




Trenchard More, John McCarthy, Marvin Minsky, Oliver Selfridge, Ray Solomonoff

AI



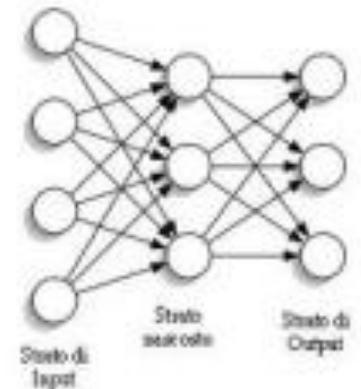
- L'AI è una disciplina scientifica che mira a definire e sviluppare programmi o macchine (software e/o hardware) che mostrano un comportamento che verrebbe definito intelligente se fosse esibito da un essere umano
- AI non è science fiction: è già nelle nostre vite ... nei sistemi di riconoscimento facciale, pagamento, traduzione, navigazione, etc.
- Come può una macchina essere considerata intelligente?
 - Cos'è l'intelligenza?
 - Approccio operativo del Test di Turing per definire l'intelligenza come imitazione basandosi sul concetto di imitazione: se una macchina fosse così abile da imitare l'uomo si potrebbe dire che pensa
- Come insegnare a una macchina a risolvere un problema
 1. **Ragionamento logico** – diciamo quali sono i passi che devono essere fatti per risolvere il problema
 2. **Machine learning** – forniamo esempi di soluzione di problemi e metodi di generalizzazione



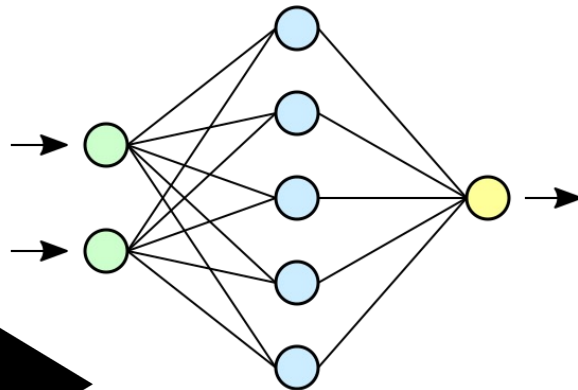
Francesca Rossi. Il confine del futuro. Possiamo fidarci dell'Intelligenza Artificiale? Feltrinelli, 2019

Bias in AI

- Biased Machine Learning?
 - Sistemi di ML per la loro natura intrinseca, possono sussumere i bias diffusi nella società su genere ed etnia, rappresentati nei data-set usati per l'addestramento, e possono portare a decisioni "unfair"
 - "Fairness is the absence of any prejudice or favoritism toward an individual or a group based on their inherent or acquired characteristics"¹
 - Consapevolezza come sviluppatori
-
- 1 N.Mehrabi, F.Morstatter, N.Saxena, K.Lerman, A.Galstyan. 2019. A survey on bias and fairness in Machine Learning. <https://arxiv.org/abs/1908.09635>



- Sistemi a scatola nera
- Subsymbolic
- Data-driven



- Addestramento di GPT-2 modelli per NLP ha consumato una enorme quantità di energia
- L'energia dell'Intelligenza Artificiale.
- 'Senza energia, l'IA non può esistere. E non parliamo di una quantità trascurabile, ma di una **fame energetica** che rischia di ridefinire le priorità di interi sistemi-paese. **I data center** sono enormi "fabbriche" costituite da moltissimi computer che lavorano insieme per addestrare e far girare gli algoritmi.'
- https://ilbolive.unipd.it/it/news/scienza-ricerca/lenergia-dellintelligenza-artificiale-intreccio?utm_source=newsletterBoLive

Face recognition



Joy Buolamwini - MIT media
lab researcher –

*Facial Recognition Is Accurate,
if You're a White Guy*



FR System

Average Accuracy:



















Lighter male 99 %

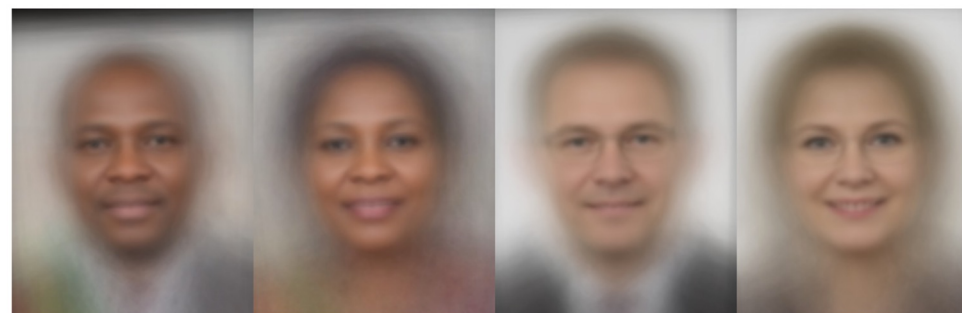
Lighter female 93 %

Darker male 88 %

Darker female 65 %

- A.I. software is only as smart as the data used to train it. If there are many more white men than black women in the system, it will be worse at identifying the black women.
- *Coded Bias*
- **Timnit Gebru**, licenziata da Google perché denunciava i pregiudizi razziali dell'Intelligenza Artificiale

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
 Microsoft	94.0% 	79.2% 	100% 	98.3% 	20.8% 
 FACE++	99.3% 	65.5% 	99.2% 	94.0% 	33.8% 
 IBM	88.0% 	65.3% 	99.7% 	92.9% 	34.4% 



Amazon Recruitment Tool

- Gli specialisti di machine-learning in Amazon hanno scoperto che “their new recruiting engine did not like women”
- Il sistema di assunzione di personale era basato su un sistema di ML
- Il sistema era stato addestrato sui cv delle persone assunte nei precedenti 10 anni. La maggior parte erano uomini quindi...
- “So the automatic recruitment tool preferred male candidates”
- E’ stato disattivato
- [Jeffrey Dastin](#). (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.



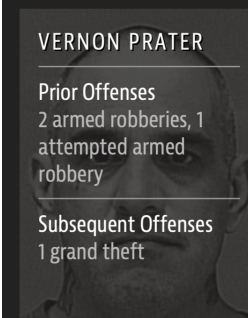
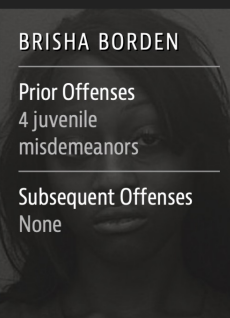
Traduttori automatici

- “Ci sono una dottoressa e un infermiere”
- Italiano/Inglese (Inglese non distingue genere)
- → “*There are a doctor and a nurse.*”
- Inglese/Italiano
- → “*Ci sono un medico e un'infermiera.*”
- Non distinguendo l'inglese tra il genere, il sistema ha dovuto ‘indovinare’ il sesso del dottore e dell'infermiere



Sistema Compas per il rischio di recidiva

 DYLAN FUGETT Prior Offense 1 attempted burglary Subsequent Offenses 3 drug possessions LOW RISK 3	 BERNARD PARKER Prior Offense 1 resisting arrest without violence Subsequent Offenses None HIGH RISK 10
---	--

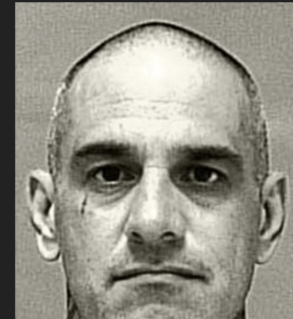
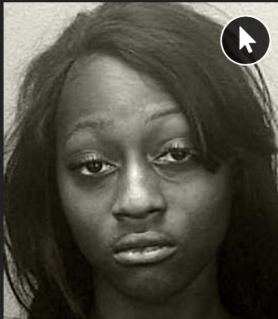
 VERNON PRATER Prior Offenses 2 armed robberies, 1 attempted armed robbery Subsequent Offenses 1 grand theft LOW RISK 3	 BRISHA BORDEN Prior Offenses 4 juvenile misdemeanors Subsequent Offenses None HIGH RISK 8
--	---

Two Drug Possession Arrests

 DYLAN FUGETT LOW RISK 3	 BERNARD PARKER HIGH RISK 10
--	--

Fugett was rated low risk after being arrested with cocaine and marijuana. He was arrested three times on drug charges after that.

Two Petty Theft Arrests

 VERNON PRATER LOW RISK 3	 BRISHA BORDEN HIGH RISK 8
---	--

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

ChatGPT

- **ChatGPT** è un modello di chatbot basato su apprendimento automatico sviluppato da OpenAI specializzato nella conversazione con un utente umano
- Poggia il suo funzionamento sul **modello linguistico**
- **GPT** sta per Generative Pre-trained Transformer

Da chatGPT3

- L'intelligenza artificiale (IA) è un ramo dell'informatica e della scienza dell'informazione che si occupa di creare macchine in grado di eseguire compiti che normalmente richiedono l'intelligenza umana. L'obiettivo principale dell'IA è quello di creare programmi o dispositivi che possono imparare, ragionare, risolvere problemi, comprendere il linguaggio naturale e riconoscere immagini e suoni, tra le altre cose.
- Ci sono diverse tecniche utilizzate nell'IA, tra cui il machine learning, la logica fuzzy, le reti neurali e l'apprendimento profondo. L'IA viene utilizzata in molti campi, tra cui la robotica, la medicina, la finanza, l'automazione industriale e i servizi di assistenza clienti.

L'IA continua a evolversi rapidamente e molti esperti ritengono che avrà un impatto significativo sulla società in futuro, portando a nuove opportunità e sfide in vari settori.



LLM e linguaggio

- Quale ruolo stanno avendo, e sempre più avranno, i modelli generativi del linguaggio come GPT, Claude o Gemini, nella promozione di un linguaggio rispettoso verso le diversità di genere?
- **Studio Unesco 2024** mostra che la Generative AI ha rilevato evidenze di stereotipi di genere, di etnia, ect. Le donne sono spesso associate a ruoli domestici mentre gli uomini a ruoli di business e carriera.
<https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>



Scopri lo stereotipo (2)



ChatGPT: Immagine di personale medico



ChatGPT: Immagine di personale di segreteria

Source: slide T. Catarci, PD, 10.03.25

S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 16 maggio 2025

L'AI mente non sapendo di mentire¹

- Può soffrire di “allucinazioni”. Ovvero «il modello può affermare con sicurezza cose/fatti che in realtà sono totalmente inventati»
- ‘Uno dei principali vantaggi della Chat GPT è la sua capacità di “imparare” dalle conversazioni che ha con gli utenti’
- Il sistema ha appreso a non contraddire l’utente, si mostra accondiscendente con l’utente e da risposte **plausibili**. Ma la **plausibilità è diversa dalla verità**. Di qui nasce un forte pericolo. Il sistema può generare contenuti falsi, inappropriati e pericolosi. Tali contenuti poi vengono divulgati a un pubblico ampio creando false credenze. E’ necessario un controllo attento per evitare diffusione di falsità e possibilità di discriminazione a livello sociale.



¹ Francesca Rossi. Intelligenza Artificiale. Come funziona e dove
Ci porta la tecnologia che sta trasformando il mondo. Laterza 2024

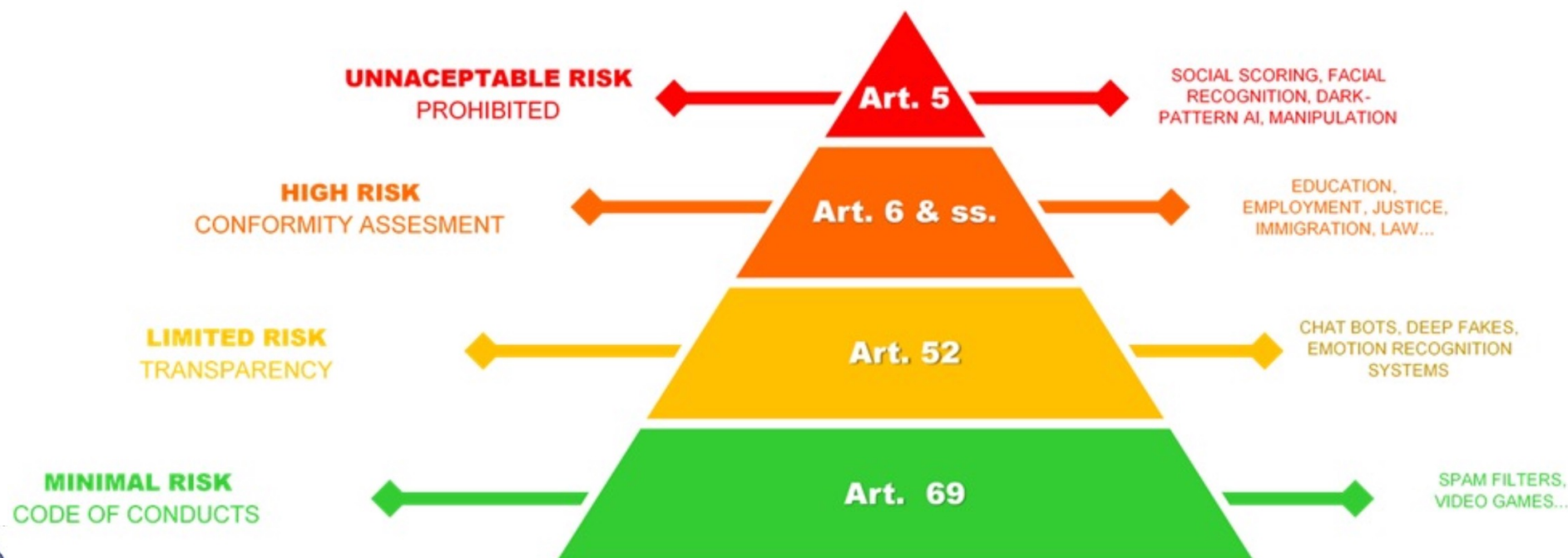
Questioni di genere?

- Interrogato su questioni di genere e se è vero che l'AI è ancora nelle mani degli uomini
- **«Più donne devono progettare i sistemi di AI, altrimenti si perpetuano i pregiudizi»**

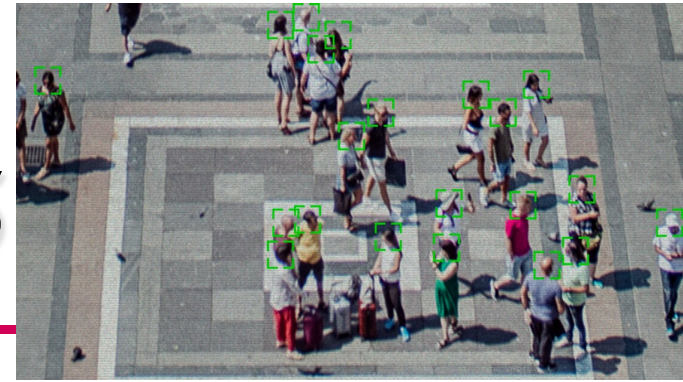


AI-Act - A EU pyramid of ranking

Levels of risk



AI-Act articolo 5



- Sistemi di IA per l'identificazione biometrica a distanza in tempo reale in spazi accessibili al pubblico
- Sistemi per la sorveglianza massiva indiscriminata: uso FR in spazi pubblici
- Sistemi di IA che manipolano il comportamento umano
- Sistemi per il riconoscimento delle emozioni sul luogo di lavoro o nelle istituzioni scolastiche
- Sistemi di social scoring
- Sistemi per le pratiche di polizia predittiva



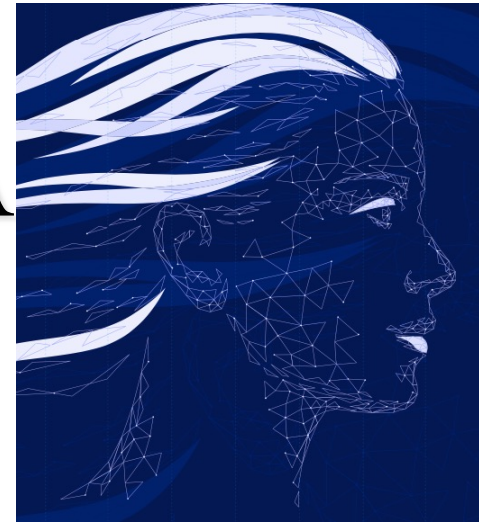
Possiamo fidarci?

- Le tecniche di IA che vengono utilizzate da gran parte delle applicazioni sono degne della nostra fiducia (“trustworthy”)?
- Non possiamo permettere a una tecnologia e al suo uso di danneggiare valori umani fondamentali
- AI trasparente, spiegabile, rispettosa della privacy e scevra da pregiudizi
- Promuovere la carriera delle donne in AI: solo 16% di persone high-skilled AI sono donne
- AI pensa da uomo perchè è stata pensata quasi esclusivamente da uomini (bianchi)
- L’Intelligenza Artificiale è una **faccenda da uomini bianchi**: in ambito accademico e nelle industrie le donne sono pochissime ed è esigua la presenza di neri e ispanici
- Secondo Global Gender Gap Index del WEF ci vorranno 134 anni per colmare il divario di genere



AI for good – Etica dell'AI

Trustworthy AI



- 3 prospettive:
 - **Promuovere la carriera delle donne in AI**
 - **Debiasing i data-set, lavorare su dati e algoritmi**
 - **Sviluppare formazione e disseminare una cultura di genere**
- **Saperi di genere ed etica nell'AI** - 6 crediti liberi - a Ingegneria Informatica ma aperto a tutti/e. - Silvana Badaloni e Antonio Roda' – aa 2024/25
- <https://stem.elearning.unipd.it/course/view.php?id=11253>
- Questo insegnamento mira a fornire agli studenti e alle studentesse gli strumenti critici per analizzare i rischi derivanti dalle applicazioni dell'Intelligenza Artificiale e a ideare soluzioni per progettare macchine intelligenti di cui ci si possa fidare.

Il vero scopo non è migliorare la tecnologia, ma migliorare noi stessi e la nostra società tramite l'uso della tecnologia

Francesca Rossi, IBM

S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 16 maggio 2025

Grazie!



- <https://www.youtube.com/watch?v=LQM8hqlfKO0>
- <http://www.donnescienza.it>
- <http://www.epws.org>
- <http://eige.europa.eu>
- <http://genderedinnovations.stanford.edu/case-studies/>
- <http://www.unipd.it/bilancio-genere>

L'unico modo per arrivare a una visione più ampia è essere in un punto particolare
(Donna Haraway, 1995)

S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 16 maggio 2025

References

- S. Badaloni and L. Perini (Eds) A model for building a Gender Equality Index for academic Institutions. Padova University Press 2016 <http://www.padovauniversitypress.it/publications/9788869380983>
- S. Badaloni, F. A. Lisi (2020) Towards a Gendered Innovation in AI In: G. Vizzari, et al (Eds): Proceedings of the AIXIA 2020 Discussion Papers Workshop co-located with the the 19th International Conference of the Italian Association for Artificial Intelligence (AIXIA2020), Anywhere, November 27th, 2020 <http://ceur-ws.org/Vol-2776/paper-2.pdf>
- S.Badaloni, A.Rodà (2022). Gender Knowledge and Artificial Intelligence, *1st Workshop on Bias, Ethical AI, Explainability and the role of Logic and Logic Programming, BEWARE-22, co-located with AIXIA 2022, University of Udine, Udine, Italy.* <https://ceur-ws.org/Vol-3319/paper12.pdf>
- S. Lohr (2018) Facial Recognition Is Accurate, if You're a White Guy, The New York Times.
- N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, Aram Galstyan: A Survey on Bias and Fairness in Machine Learning. CoRR abs/1908.09635 (2019)
- F. Rossi. Il confine del futuro. Possiamo fidarci dell'Intelligenza Artificiale? Feltrinelli, 2019
- [European Commission 2020] European Commission (2020). Gendered Innovation 2: How Inclusive Analysis Contributes to Research and Innovation. ISBN 978-92-76-16416-62 https://ec.europa.eu/info/news/gendered-innovations-2-2020-nov-24_en
- Ethics guidelines for trustworthy AI (2019) <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>
- I FOCUS DEL RAPPORTO ANVUR 2023 – Analisi di genere <https://www.anvur.it/wp-content/uploads/2024/01/Focus-equilibrio-di-genere-2023.pdf>
- UNESCO Report 2020, Artificial Intelligence and Gender Equality. <https://en.unesco.org/AI-and-GE-2020>
- UNESCO, OECD and IDB (2022). The effects of Artificial Intelligence on the working lives of women. <https://oecd.ai/en/women-event-2022>



Addio Giulia!

